DOI: 10.4274/jbuch.galenos.2026.50470
J Dr Behcet Uz Child Hosp 2026;16(2):127-139

Original Article

Dental Caries Management Patterns in Pediatric Patients with Chronic Medical Conditions: A Natural Language Processing Approach

Kronik Tıbbi Durumları Olan Pediyatrik Hastalarda Diş Çürüğü Yönetim Örüntüleri: Doğal Dil İşleme Yaklaşımı

Hüseyin Demir[1], Elif Ece Kalaoğlu[2], Süleyman Mertoğlu[3]

[1]İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye
[2]İstanbul Gelişim University Faculty of Dentistry, Department of Pediatric Dentistry, İstanbul, Türkiye
[3]İzmir Provincial Directorate of Health, İzmir, Türkiye

ABSTRACT

Objective: Chronic childhood diseases present significant challenges for dental caries management, yet the intellectual structure of the research domain remains poorly mapped. This study systematically characterizes the knowledge structure, thematic evolution, and citation dynamics at the intersection of pediatric chronic diseases and dental caries management.

Method: A total of 545 abstracts from the Web of Science (1990-2026) were analyzed using natural language processing. K-means clustering and principal component analysis identified intellectual groupings while Latent Dirichlet allocation enabled topic modeling. Temporal evolution and citation impact were assessed, and co-occurrence network analysis revealed conceptual linkages.

Results: Three distinct knowledge clusters emerged: Cluster 0 (specialized studies, average number of citations =27), Cluster 1 (mainstream paradigm, average number of citations =80), and Cluster 2 (emerging research, average number of citations =19). Five optimal research themes were identified (coherence =0.365): General pediatric oral health epidemiology, childhood caries risk factors, patient-centered disease-dental health relationships, comparative methodological studies, and asthma-focused research. Temporal analysis revealed four phases: latent (1990-2000), awakening (2000-2006), initial growth (2006-2015), and accelerated expansion (post-2020). Patient-centered topics peaked at 18-19 publications annually by 2026. Disease-specific analysis showed concentration on diabetes (124 publications), asthma (82), and cerebral palsy (59), with notable underrepresentation of sickle cell anemia and juvenile arthritis. Network analysis identified "oral health," "disease," and "caries" as central nodes, linking clinical outcomes to quality-of-life concepts.

Conclusion: The domain has shifted from descriptive epidemiology toward patient-centered and community-oriented approaches. Persistent thematic and disease-specific gaps highlight the need for integrative, interdisciplinary research addressing neglected chronic pediatric conditions.

Keywords: Dental caries, chronic diseases, children, natural language processing, topic modeling

ÖZ

Amaç: Kronik çocukluk çağı hastalıkları, diş çürüğü yönetimi açısından önemli zorluklar ortaya koymaktadır. Ancak araştırma alanının tematik yapısı yeterince araştırılmamıştır. Bu çalışma, pediyatrik kronik hastalıklar ile diş çürüğü alanının bilgi yapısını, tematik evrimi ve atıf dinamiklerini sistematik olarak araştırmıştır.

Yöntem: Web of Science'tan (1990-2026) toplam 545 özet, doğal dil işleme kullanılarak analiz edilmiştir. K-means kümeleme ve temel bileşenler analizi tematik gruplandırmada, Gizli Dirichlet ayırımı ise konu modelleme amacıyla kullanılmıştır. Zamansal evrim ve atıf etkisi değerlendirilmiş, sosyal ağ analizi ise kavramsal bağlantıları göstermiştir.

Bulgular: Analizler sonucunda üç farklı bilgi kümesi ortaya çıkmıştır: Küme 0 (uzmanlaşmış çalışmalar, ortalama atıf =27), Küme 1 (ana akım paradigma, ortalama atıf =80) ve Küme 2 (gelişmekte olan araştırma, ortalama atıf =19). Beş araştırma teması belirlenmiştir (tutarlılık =0,365): Genel pediyatrik ağız sağlığı epidemiyolojisi, çocukluk çağı çürük risk faktörleri, hasta merkezli hastalık-diş sağlık ilişkileri, karşılaştırmalı

Received: 04.02.2026
Accepted: 10.06.2026
Publication Date: 21.08.2026

Corresponding Author
Hüseyin Demir,
İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye
E-mail: huseyin.demir@ikc.edu.tr
ORCID: 0000-0002-8990-7228

Cite as: Demir H, Kalaoğlu EE, Mertoğlu S. Dental caries management patterns in pediatric patients with chronic medical conditions: a natural language processing approach. J Dr Behcet Uz Child Hosp. 2026;16(2):127-139

metodolojik çalışmalar ve astım odaklı araştırma. Zamansal analiz dört aşama ortaya koymuştur: gizil dönem (1990-2000), uyanış (2000-2006), ilk büyüme (2006-2015) ve hızlandırılmış genişleme (2020 sonrası). Hasta merkezli konular 2026 yılına kadar yılda 18-19 yayına ulaşarak zirve yapmıştır. Hastalığa özgü analiz, diyabet (124 yayın), astım (82) ve serebral palsi (59) üzerinde yoğunlaşırken, orak hücreli anemi ve juvenil artrit yetersiz temsil edilmiştir. Sosyal ağ analizi, klinik sonuçları yaşam kalitesi kavramlarına bağlayan merkezi düğümler olarak yer almıştır.

Sonuç: Alan, tanımlayıcı epidemiyolojiden hasta merkezli ve toplum odaklı yaklaşımlara doğru kaymıştır. Kalıcı tematik ve hastalığa özgü boşluklar, ihmal edilmiş kronik pediyatrik hastalıkları ele alan bütünleştirici, disiplinler arası araştırma ihtiyacını vurgulamaktadır.

Anahtar kelimeler: Diş çürüğü, kronik hastalıklar, çocuklar, doğal dil işleme, konu modellemesi

INTRODUCTION

Dental caries is among the most prevalent chronic diseases of childhood, affecting more than 510 million children worldwide through decay of the primary dentition[1]. This burden is particularly pronounced in children living with chronic systemic diseases, where complex, bidirectional interactions between the underlying medical condition(s) and oral health often exacerbate disease severity[2-4]. Children affected by conditions such as cerebral palsy, asthma, diabetes mellitus, cardiovascular anomalies, cystic fibrosis, and neurodevelopmental disorders consistently experience marked oral health disparities compared with their typically developing peers. These disparities commonly present as higher rates of dental caries, periodontal disease, developmental enamel defects, and a reduced oral health-related quality of life[5-9]. In cerebral palsy, for instance, motor limitations and orofacial dysfunction can make daily oral hygiene difficult to maintain, increasing susceptibility to dental disease(s)[2]. Similarly, children with asthma are exposed to a more cariogenic oral environment as a result of long-term inhaled corticosteroid use, habitual mouth breathing, and reduced salivary flow (hyposalivation)[8]. Children with chronic kidney disease also face oral health challenges, as salivary hypofunction and biochemical alterations disrupt normal oral homeostasis and relevant protective mechanisms[7].

The therapeutic implications of dental caries in children with medical complexities extend well beyond disease prevalence. These patients often require invasive restorative procedures, including pulp therapy, preformed metal crowns, or tooth extractions. In young children or those with behavioral and cognitive challenges, such interventions frequently necessitate pharmacological support through conscious sedation or general anesthesia[10,11]. Delivering anesthesia to children with chronic systemic diseases is particularly demanding due to preexisting medical instability, airway and anatomical variations, cardiorespiratory compromise, and the need for multidisciplinary preoperative assessment. As a result, treatment planning is often constrained by medical contraindications, limited access to hospital-based dental services, and heightened safety requirements, which may lead to delayed care, compromised treatment decisions, or unnecessarily aggressive interventions[12,13].

These clinical challenges are reflected in the current research landscape, which remains notably fragmented. Much of the existing literature focuses on individual chronic conditions in isolation, with limited integration of findings into broader, cross-cutting frameworks for managing dental caries in pediatric populations with medical complexities[2-4,6-8,14]. This compartmentalized approach restricts synthesis of interdisciplinary knowledge and hinders the development of unified clinical guidelines and health policies. Although individual studies provide valuable insights, still scarce number of comprehensive analyses mapped the intellectual structure and temporal evolution of this domain.

Advances in natural language processing (NLP) and bibliometric methods offer powerful tools for systematically synthesizing large bodies of medical literature[15]. By enabling automated extraction of themes, identification of research clusters, and visualization of scholarly relationships, these approaches help uncover emerging trends, shifts in research focus, and underexplored areas within complex domains[16]. Applying such methods to pediatric dental research allows for a more integrated understanding of how knowledge has developed and where critical gaps persist. Despite the growing body of literature on pediatric oral health, the intellectual structure of research on dental caries management in children with chronic diseases remains poorly characterized. It is still unclear which thematic clusters dominate this field, how research priorities have evolved over time, and where critical knowledge gaps persist particularly at the intersection of specific chronic conditions and caries prevention or treatment strategies. Traditional narrative reviews have addressed isolated aspects of this topic but fall short of providing a comprehensive, data-driven mapping of the field's conceptual landscape. To address these gaps, the present study employed an NLP-based analytical framework to systematically map the intellectual structure, thematic evolution, and emerging research priorities in dental caries management among children with chronic diseases.

MATERIALS and METHODS

This study applied an integrated NLP framework to analyze the thematic structure, temporal evolution, and conceptual relationships in the literature on dental caries management in children with chronic diseases. By combining text preprocessing, machine learning-based clustering, topic modeling, temporal dynamics and network analysis, the methodology enables a comprehensive and data-driven exploration of research patterns and trends in this domain.

Data Collection and Preprocessing

This study analyzed 545 peer-reviewed scientific publications retrieved from the Web of Science (WoS) database on January 15, 2026 concerning management of dental caries in children with chronic diseases. The literature search was conducted using a comprehensive strategy combining four main conceptual blocks: chronic diseases, pediatric populations, dental caries, and management or preventive approaches. The following search queries were applied:

("chronic disease" OR "chronic illness" OR "chronic condition" OR "children with special health care needs" OR "diabetes" OR "asthma" OR "epilepsy" OR "congenital heart disease" OR "cerebral palsy" OR "chronic kidney disease" OR "cystic fibrosis") AND ("child" OR "children" OR "pediatric" OR "adolescent") AND ("dental caries" OR "tooth decay" OR "early childhood caries") AND ("management" OR "prevention" OR "treatment" OR "oral health care" OR "dental care" OR "caries management" OR "fluoride therapy" OR "fissure sealants" OR "preventive dentistry" OR "minimally invasive dentistry" OR "behavior management" OR "special care dentistry" OR "oral health" OR "pediatric dentistry").

The full search string included terms such as chronic disease, diabetes, asthma, child, pediatric, dental caries, and oral health management, and was limited to English-language articles indexed in Science Citation Index Expanded, Social Sciences Citation Index, and Emerging Sources Citation Index (ESCI) databases, in accordance with the preferred reporting items for systematic reviews and meta-analyses (PRISMA 2020) guidelines[17]. Peer-reviewed original research articles focusing on management, prevention, or treatment of dental caries in pediatric populations with chronic diseases were analytically reviewed. Relevant studies involving children and adolescents, published in English, and providing empirical data were considered eligible. Review articles, systematic reviews, meta-analyses, editorials, letters, conference abstracts, non-peer-reviewed publications, studies not directly addressing dental caries or not involving children with chronic conditions were excluded. The study selection process is illustrated in Figure 1.

A total of 1,389 records were identified through the WoS database. After the removal of 187 duplicates,

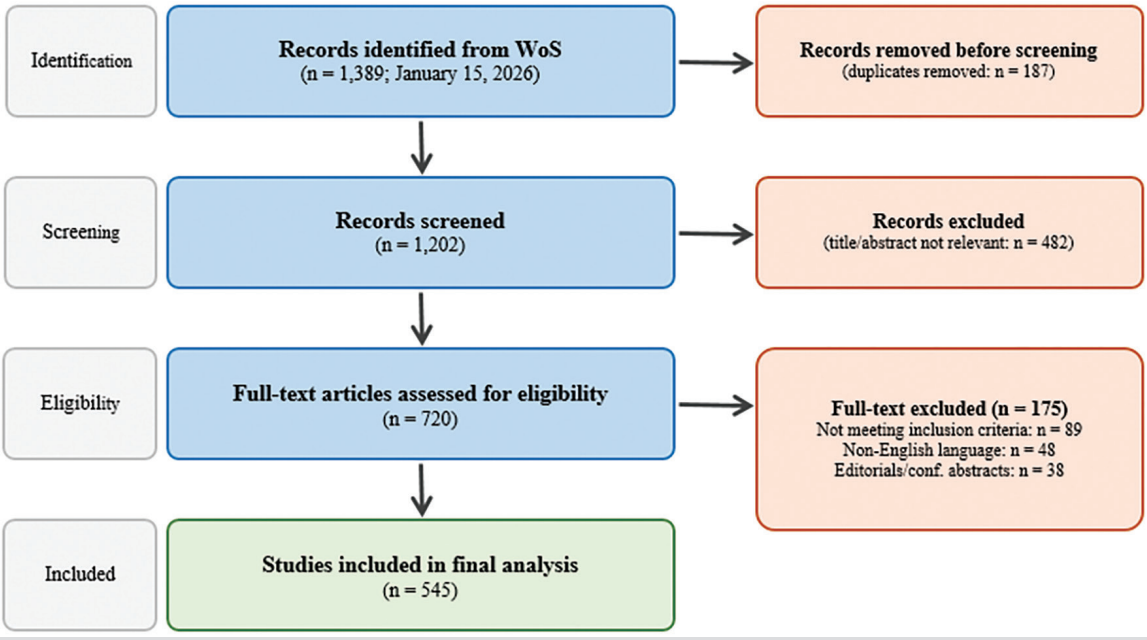

Figure 1. PRISMA flow diagram of the study selection process

PRISMA: Preferred reporting items for systematic reviews and meta-analyses, WoS: Web of Science

1,202 records remained for screening, of which 482 were excluded based on title and abstract. Subsequently, 720 full-text articles were assessed for eligibility, and 175 of them were excluded from the analyses due to not meeting the inclusion criteria i.e. articles published in non-English languages, or being editorials or conference abstracts. Ultimately, 545 studies were included in the final analysis.

The final dataset comprised three core variables: abstract, citation count, and publication year. Text preprocessing was performed exclusively on the abstract. The preprocessing pipeline consisted of: (1) conversion to lowercase for standardization, (2) removal of punctuation marks and numerical expressions using regular expressions, (3) elimination of English stop-words using the Natural Library Toolkit (NLTK) library, and (4) lemmatization using the WordNetLemmatizer to reduce words to their root forms. The cleaned and standardized texts were stored in a new variable, forming the textual input for subsequent NLP and modeling analyses[18-20].

Semantic Analysis with Word2Vec

A Word2Vec model was trained using the Gensim library on the preprocessed abstracts to learn contextual semantic representations of words as 100-dimensional vectors[21]. Document-level numerical representations were created by calculating the arithmetic mean of word vectors within each abstract. These document vectors, stored in the document vector column, served as input for subsequent clustering and dimensionality reduction analyses[22].

Optimal Cluster Number Determination

The optimal number of clusters was determined using the Elbow method[23]. K-means clustering was performed on document vectors with varying numbers of clusters (k=2 to k=10), and inertia values (sum of squared distances from each point to its cluster center) were calculated for each k[24]. The elbow point, where the marginal decrease in inertia diminished substantially, was identified through visual inspection of the inertia plot[23]. Based on this analysis, k=3 was selected as the optimal cluster number, balancing model complexity with explanatory power.

Clustering Analysis

K-means clustering was applied to the document vectors with k=3 clusters. Citation counts were incorporated as weighting factors in the clustering process, allowing highly cited articles to exert greater influence on cluster center determination. This approach ensured that influential or impactful publications shaped the cluster structure. Each article was assigned a cluster label representing its thematic group.

Citation Analysis

For each cluster, average citation counts were calculated to assess scientific impact and visibility. This bibliometric analysis revealed differences in citation performance across clusters, indicating varying levels of influence and recognition within the scientific community. Results were visualized using bar charts to facilitate comparisons.

Optimal Topic Number Determination

To identify the optimal number of topics for LDA (Latent Dirichlet Allocation) modeling[25] coherence score analysis was conducted[26]. Using Gensim's CoherenceModel, coherence scores were calculated for topic numbers ranging from 2 to 10. Coherence scores measure the semantic similarity between high-scoring words in topics, with higher values indicating more interpretable and meaningful themes[26]. The coherence scores were plotted, and the optimal number of topics yielding the highest coherence score was selected. Based on this analysis, 5 topics were determined to provide the best thematic representation of the literature.

LDA Topic Modeling

LDA topic modeling was implemented using processed abstracts to identify thematic structures in the literature[25]. A dictionary and document-term matrix were constructed using Gensim's Dictionary and corpus functions[27]. The LDA model was trained with 5 topics (as determined by coherence analysis), extracting the most prominent keywords and their weights for each topic. Topic distributions were calculated for each document, and the dominant topic and its probability were assigned to each article.

Cluster Visualization with Principal Component Analysis PCA

To visualize clustering results, document vectors were reduced to two dimensions using PCA. Technique PCA1 (x-axis) represented the dimension explaining the most variance, while PCA2 (y-axis) captured the second most important variance source. The two-dimensional projections were plotted as scatter plots, with points colored by cluster labels, enabling visual assessment of cluster separation and spatial relationships.

Temporal Dynamics of Clusters and Topics

Temporal trends of the three clusters were analyzed by grouping articles by their publication years. The number

of articles in each cluster was calculated annually from 1990 to 2026. These trends were visualized using line plots with Seaborn[28], revealing how cluster popularity evolved over time and identifying periods of growth, stability, or decline for different research paradigms. Similarly, temporal trends of the five LDA topics were examined by grouping articles by year and dominant topic. Annual article counts for each topic were calculated and visualized using line plots. This analysis revealed the evolution of research themes over time, identifying emerging, stable, and declining topics within the literature.

Topic Word Clouds Analysis

Word clouds were generated for each of the 5 LDA topics using the WordCloud library[29]. Keywords were sized proportionally to their probability weights within each topic, providing intuitive visual representations of thematic content. These visualizations facilitated rapid interpretation of topic characteristics and complemented the quantitative topic modeling results.

Chronic Disease Content Analysis

Keywords related to common chronic pediatric diseases (diabetes, asthma, obesity, cerebral palsy, congenital heart disease, cystic fibrosis, epilepsy, autism, sickle cell anemia, juvenile arthritis) were defined and searched within the processed abstracts. The frequency and percentage of articles mentioning each condition were calculated using named entity recognition model[30], revealing the chronic diseases that were most prominently studied in relation to management of dental caries. Results were ranked in descending order and visualized using bar charts.

Social Network Analysis

A social network was constructed to examine relationships among key concepts related to the management of dental caries, health outcomes, and complications[31]. Keywords representing management strategies (management, prevention, treatment, intervention), health outcomes (oral health, quality of life, outcome, wellbeing), and clinical complications (pain, abscess, disease, complication, infection, inflammation) were identified. A co-occurrence matrix was created by calculating how frequently each keyword pair appeared together in abstracts. Using NetworkX package for the Python programming language, this matrix was transformed into a network graph where nodes represented keywords and edges represented frequencies of co-occurrences, with edge thickness proportional to co-occurrence strength. The network was visualized using Matplotlib library[32] to reveal conceptual relationships and thematic clusters within the literature.

Statistical Analysis

All computational analyses were performed using Google Colaboratory[33]. Python platform was used as the primary programming language throughout the analysis process[34]. The computational workflow was supported by various libraries commonly used in scientific research. These include NumPy[35] for numerical operations, Pandas[36] for data processing and table analyses, Matplotlib[32] and Seaborn[28] for visualization, Scikit-learn[37] for ML operations, Statsmodels[38] for time series modeling, and WordCloud[29] for word cloud visualizations.

To enhance language quality, AI tools were utilized during manuscript preparation. Specifically, OpenAI's ChatGPT[39] facilitated English translation, and Anthropic's Claude AI[40] was employed for linguistic review and stylistic refinement.

RESULTS

This study employed an integrated NLP framework to examine the thematic structure, temporal evolution, and conceptual relationships in the literature on the management of dental caries among children with chronic diseases. Through the combined use of text preprocessing, machine learning-based clustering, topic modeling, temporal analysis, and network analysis, the findings provide a comprehensive, data-driven overview of research patterns and emerging trends in this domain. The findings of optimal number of clusters, clustering analysis, citation performance, and coherence scores are presented in Figure 2.

At k=2, the inertia value reached approximately 19,600, indicating insufficient granularity and excessive within-cluster variance (Figure 2). A dramatic reduction to approximately 15,000 was observed at k=3, representing a marginal improvement of 4,600 units-the most substantial decrease across all iterations. This inflection point constituted a distinct elbow, beyond which incremental cluster additions yielded diminishing returns, primarily fragmenting existing structures rather than revealing novel patterns. Consequently, 3 clusters were determined as optimal for subsequent analyses.

PCA reduced the multidimensional feature space to two principal components (PC1 and PC2), capturing the dominant variance axes within the corpus. Cluster 0 (turquoise; n≈50-70) exhibited spatial isolation on the negative PC1 axis, demonstrating tight cohesion

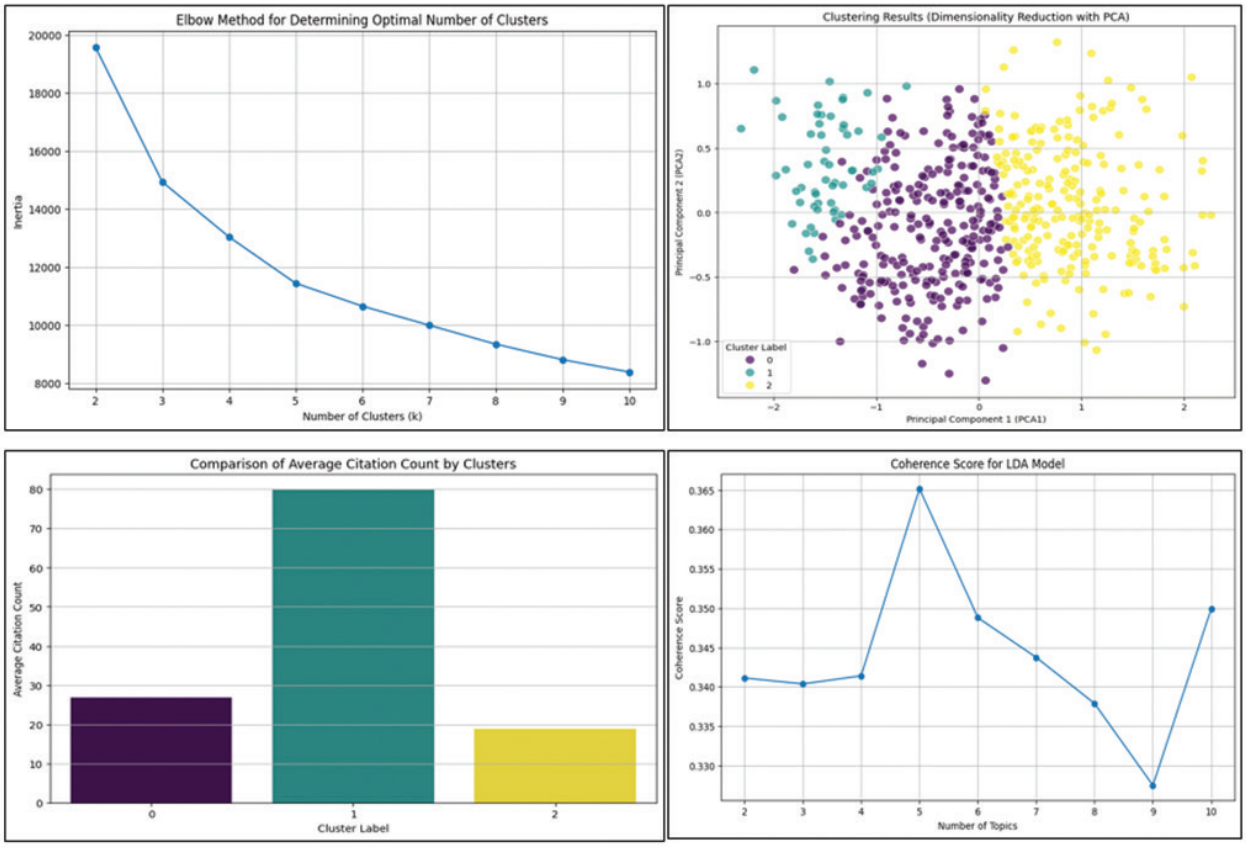

Figure 2. Optimal number of clusters, clustering analysis, citation performance, and coherence scores

indicative of methodological homogeneity and potentially representing a specialized research niche or disease-specific focus. Cluster 1 (purple; n>200) occupied the central region with substantial overlap zones, suggesting its role as the disciplinary mainstream and a transitional bridge between research paradigms. Cluster 2 (yellow; n≈150) displayed greater dispersion along the positive PC1 axis, potentially reflecting thematic heterogeneity or emerging research trajectories. The pronounced separation along PC1 indicated this component as the primary discriminator between clusters, while boundary overlaps suggested continuous epistemological transitions rather than discrete categorical divisions.

Citation performance demonstrated marked heterogeneity across clusters. Cluster 1 achieved the highest mean citation count (M=80), consistent with its central positioning and substantial size, indicative of an established, well-integrated research paradigm. Cluster 0 exhibited moderate impact (M=27), approximately one-third that of Cluster 1, suggesting specialized but recognized scholarly contributions. Cluster 2 recorded the lowest mean citations (M=19), potentially reflecting its emergent nature or peripheral positioning within the intellectual structure of the domain. These differential citation patterns underscore the stratified maturity levels across research clusters: Cluster 1 representing canonical knowledge, Cluster 0 denoting specialized expertise, and Cluster 2 signifying nascent or exploratory research domains.

Coherence score analysis across 2-10 topics identified the optimal LDA configuration. The coherence metric peaked at 0.365 for k=5 topics, representing the optimal balance between thematic granularity and interpretability. Coherence degradation beyond five topics suggested model overfitting and excessive fragmentation, while lower values indicated inadequate capture of the corpus's thematic complexity. The 5-topic model was therefore adopted for subsequent semantic analysis. The findings related to the temporal dynamics of the clusters are presented in Figure 3.

Longitudinal analysis revealed divergent temporal trajectories across clusters. Cluster 0 demonstrated sustained growth post-2005, with pronounced acceleration after 2020, culminating in peak activity in

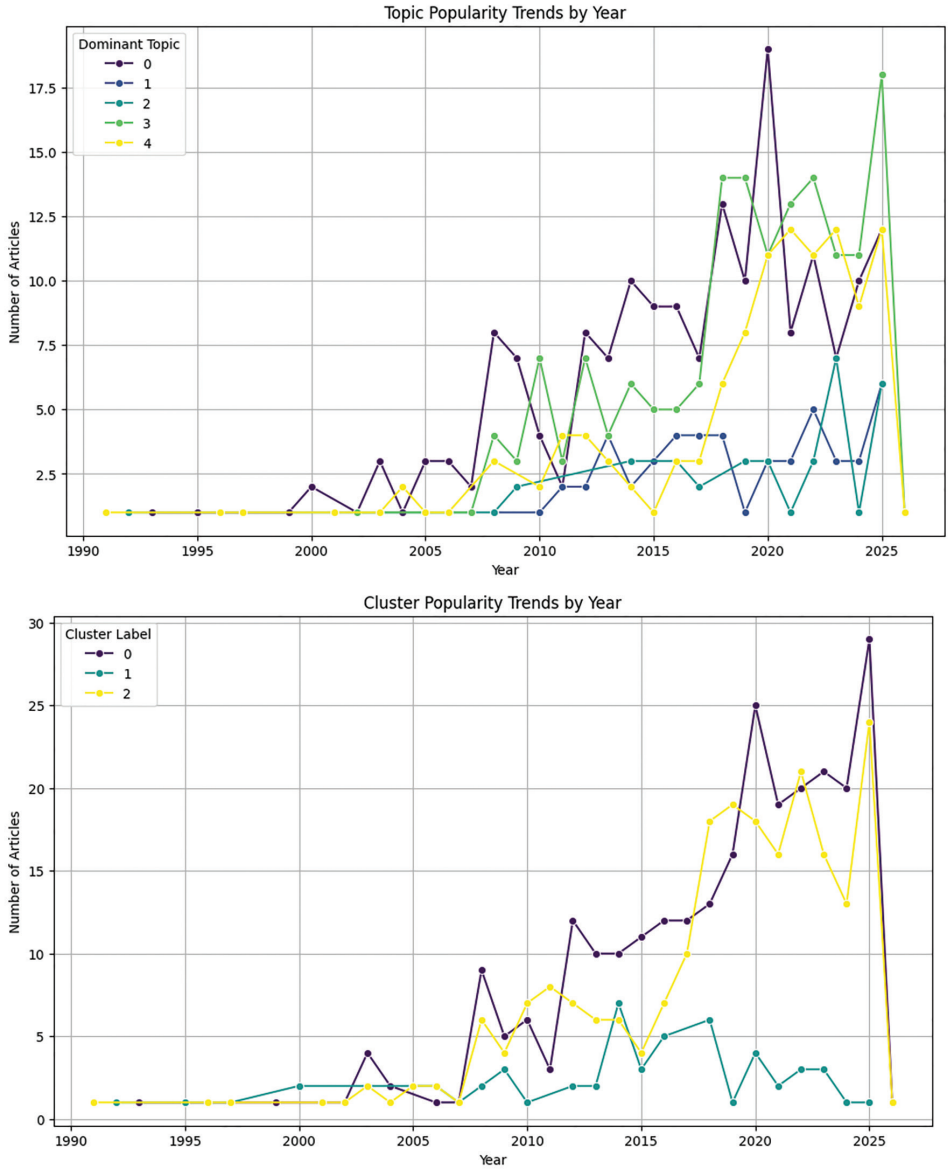

Figure 3. Temporal dynamics of clusters and topics

2026. Cluster 2 exhibited episodic expansion with distinct peaks in 2018 and 2026, suggesting periodic surges in scholarly attention. Conversely, Cluster 1 showed modest growth followed by substantial decline post-2020, potentially indicating paradigmatic saturation or domain reorientation. The temporal distribution of LDA topics spanning 1990-2026 revealed distinct evolutionary phases. The pre-2000 period constituted a latent phase characterized by minimal publication activity. Progressive intensification commenced post-2000, with Topic 0 and Topic 2 demonstrating particularly robust growth after 2006 and accelerated expansion during 2015-2020. Post-2020, Topic 0 emerged as the dominant research theme, accompanied by substantial increases in Topics 2 and 4. These patterns reflect the domain's increasing prioritization of dental caries management in pediatric populations with chronic disease(s), with accelerating research momentum in recent years. Word clouds representing the clustered terms under each topic are presented in Figure 4.

LDA-derived word clouds illuminated the semantic architecture of five distinct research themes.

Topic 0: General Pediatric Oral Health Epidemiology

Represents the domain's foundational dimension. High-frequency terms including child, health, oral, dental, and caries indicate predominance of epidemiological and descriptive investigations. Co-occurring terms (prevalence, disease, care, study, and result) suggest focus on quantification of disease burden and assessment of population-level health status.

Topic 1: Risk Factors for Childhood Dental Caries

Adopts an analytical-etiological orientation. Prominence of caries, dental, health, risk, and age reflects research examining risk determinants, age-stratified patterns, and impact of chronic diseases on the pathogenesis of caries through predictive modeling and risk stratification frameworks.

Topic 2: Patient-Centered Chronic Disease-dental Health Interface

Embodies a holistic, patient-oriented paradigm. Central positioning of patient, health, caries, dental diseases, and particularly diabetes indicates examination of chronic disease-oral health relationships through experiential, care process, and community intervention lenses. Terms as community and care underscore public health and multidisciplinary perspectives.

Topic 3: Comparative Methodology and Controlled Research

Constitutes the strongest methodological theme. Prominence of group, control, index, study, and child reflects high-evidence designs employing control groups, comparative frameworks, and standardized metrics (e.g., Decayed, Missing, Filled Teeth: DMFT index), focusing on efficacy and comparative effectiveness of the intervention.

Topic 4: Disease-specific Approaches and Asthma

Reflects condition-targeted research. The salience of asthma indicates concentrated investigation of asthma-caries relationships, while risk, dental, caries, child, and patient highlight focused analyses of specific risk populations and targeted dental management in defined chronic disease cohorts. Cross-topic analysis revealed that while core concepts (child, dental, caries, health, and oral) permeate all themes, each topic maintains distinctive keyword profiles reflecting unique research questions and methodological orientations. Temporal dynamics indicated particularly robust recent growth in Topic 2, highlighting the ascending importance of patient-centered, holistic approaches in contemporary pediatric dental research. Collectively, these five topics demonstrate the multidimensional, interdisciplinary, and evolutionary nature of dental caries management research in chronic diseases in pediatric populations. The findings related to chronic diseases highlighted in the dental health literature are presented in Figure 5.

Analysis of representation of chronic disease revealed substantial heterogeneity. Diabetes dominated with 124 publications, followed by asthma (n=82) and cerebral palsy (n=59). Obesity, congenital heart disease, cystic fibrosis, and epilepsy received moderate scholarly attention. Autism appeared in limited investigations, while sickle cell anemia and juvenile arthritis emerged as notably underrepresented conditions. This distribution pattern indicates concentrated research interest in

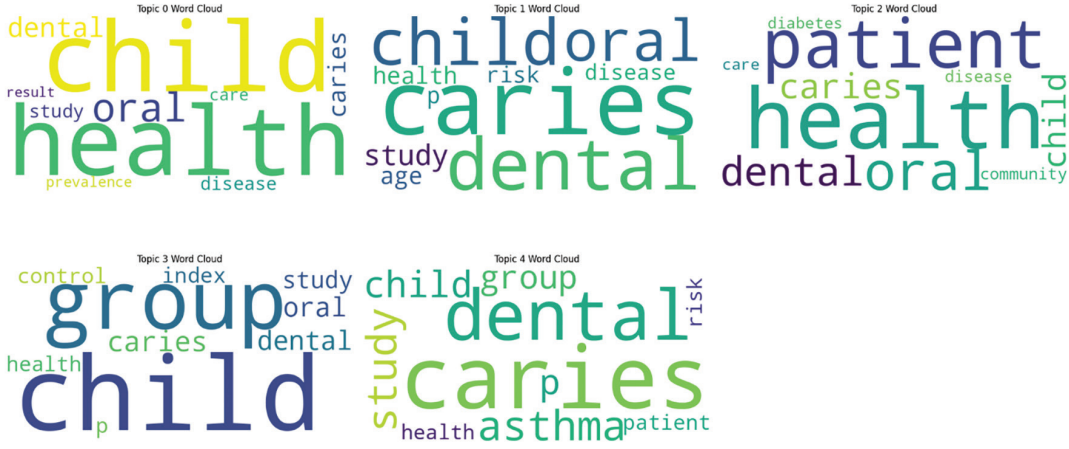

Figure 4. Word clouds of topics

specific chronic conditions, with significant knowledge gaps persisting across numerous populations with pediatric chronic diseases. The findings of the social network analysis of keyword "co-occurrence" in the pediatric dental literature are presented in Figure 6.

Social network analysis of keyword co-occurrence patterns revealed complex semantic relationships within the corpus. Network architecture featured nodes (keywords) connected by edges (co-occurrences), with node size reflecting term frequency and edge thickness indicating co-occurrence strength. The network core comprised three dominant nodes-oral health, disease, and caries-establishing the domain's conceptual foundation.

Within the clinical outcomes cluster, exceptionally strong linkage between abscess and pain reflected the clinical reality of manifestation of pain in advanced carious lesions with abscess formation. Pain functioned as a bridging construct connecting clinical symptomatology to patient-reported outcomes including impact, quality of life, and wellbeing. The proximal positioning of quality of life and wellbeing underscored the growing emphasis on patient-centered outcome assessment in pediatric research, signaling progression beyond purely clinical endpoints toward comprehensive evaluation of child wellbeing.

The management strategies in cluster projects demonstrated management as a hub connected to treatment, prevention, intervention, and outcome, representing research on therapeutic and preventive approaches to dental caries in chronic disease populations. The prominence of prevention emphasized proactive strategies (fluoride therapy, dietary modification, oral hygiene education), while intervention encompassed both clinical and educational modalities.

Regarding pathogenesis and complications, nodes including infection, inflammation, complication, and severity represented dimensions of disease process. The node status encompassed both dental status (e.g., DMFT indices) and general health status. Notably, disease burden appeared relatively isolated, suggesting limited integration of disease burden frameworks within mainstream research discourse.

Network topology analysis demonstrated non-random keyword distribution organized according to semantic proximity. Robust connections characterized the oral health-disease-caries triad and the quality of life-wellbeing dyad, while weak or absent linkages appeared between semantically distant terms (burden-abscess). The multi-centered network structure indicated that distinct conceptual clusters maintained dense internal

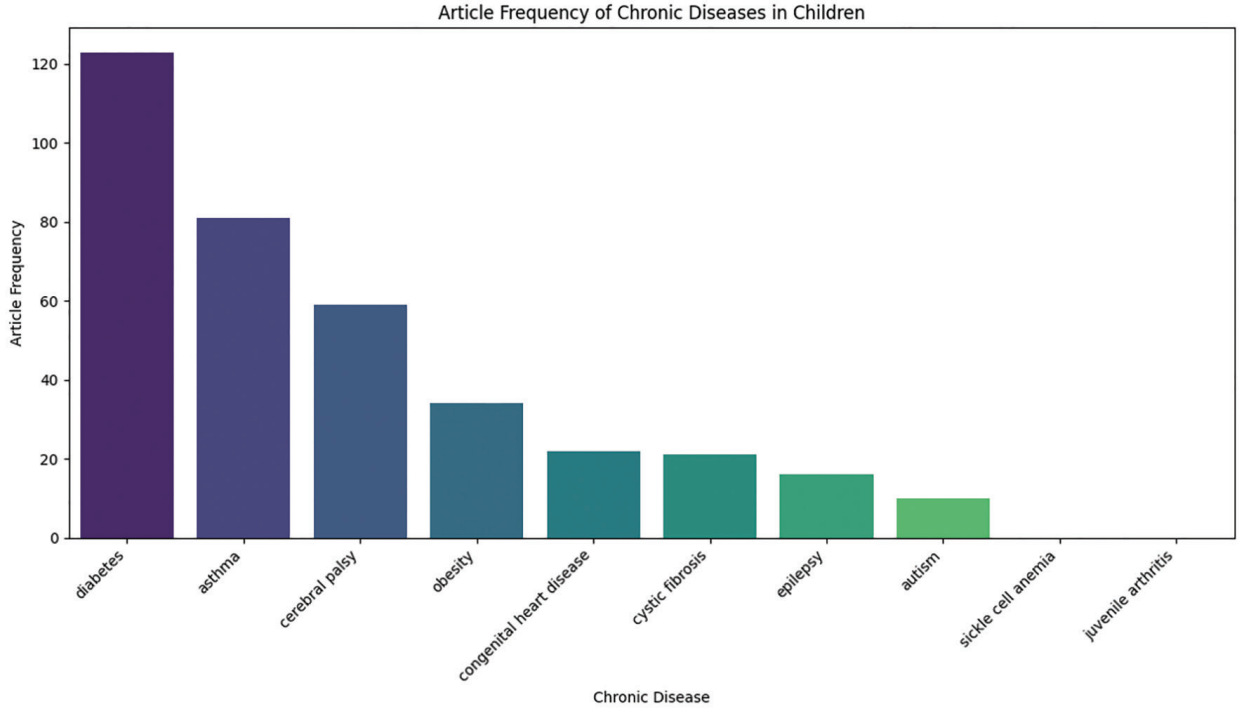

Figure 5. Chronic diseases identified in the pediatric dental health literature

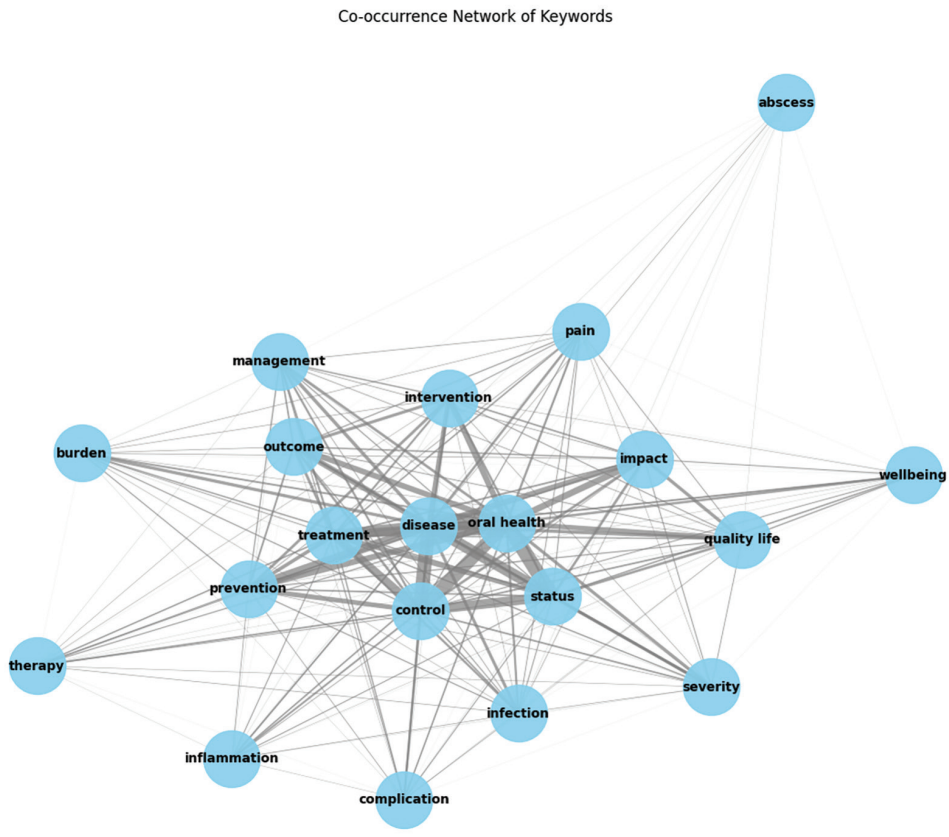

Figure 6. Social network analysis of pediatric dental literature

connectivity while bridging nodes facilitated inter-cluster integration. This complex topological architecture underscores that management of dental caries in pediatric populations with chronic disease constitutes a multifactorial, multidisciplinary, and holistic research domain requiring integrated approaches across clinical, preventive, and patient-centered dimensions.

DISCUSSION

This study provides a comprehensive overview of the scientific literature on the management of dental caries in children with chronic diseases and highlights important patterns in how this domain has evolved over time. The findings suggest that research in this area is heterogeneous, with varying thematic emphases and uneven attention across chronic disease groups. Collectively, the literature reflects both growing recognition of oral health as a component of chronic disease care and persistent gaps that may have implications for clinical practice and health equity.

The identification of three distinct research clusters suggests that the existing literature remains largely structured around disease-specific perspectives rather than integrated models of care. The body of work with the highest citation iimpact reflects the established research foundation documenting the disproportionate burden of oral disease among children with chronic conditions. Seminal studies, including that of Thikkurissy and Lal[5], consistently report higher prevalence of dental caries, periodontal disease, and developmental enamel defects in medically complex children compared with their healthy peers. The strong scholarly impact of this literature likely reflects both its clinical relevance and its methodological rigor, particularly in elucidating the bidirectional relationship between systemic disease and oral health. Previous investigations have demonstrated how chronic conditions such as cerebral palsy and metabolic disorders influence oral hygiene capacity, salivary function, inflammatory responses, and disease susceptibility, while untreated oral infections may further compromise systemic health[2-4].

Highly specialized investigations have focused on specific chronic conditions, most notably cerebral palsy and chronic kidney disease. These studies contribute important condition-specific insights, such as the effects of motor impairment and orofacial dysfunction on oral hygiene practices in children with cerebral palsy, and the role of reduction in salivary flow (hyposalivation) and biochemical alterations in increasing susceptibility to caries among children with chronic kidney disease[2,7]. Although clinically relevant, this body of work often remains embedded within subspecialty frameworks, which may limit its integration into broader pediatric oral health strategies and interdisciplinary care models.

In contrast, a growing body of literature reflects a shift toward patient-centered and integrated approaches to oral health care in children with chronic diseases. While this research has not yet achieved the same level of scholarly visibility as more established studies, its marked expansion after 2020 suggests increasing awareness that effective caries management must extend beyond disease-specific outcomes. Recent studies have emphasized quality of life, functional capacity, and psychosocial wellbeing as integral components of oral health, particularly for children living with complex medical conditions[6]. The relatively limited citation impact of this emerging literature likely reflects the inherent delay between conceptual innovation and its widespread incorporation into clinical practice. It is also worth noting that the comparatively lower citation counts observed for publications from 2021 onward should be interpreted with caution. While there is a natural lag between conceptual innovation and clinical integration, it must equally be recognized that more recently published studies have had substantially less time to accumulate citations relative to earlier works. This temporal citation bias is an inherent characteristic of bibliometric analyses and does not necessarily reflect the actual scientific impact or clinical relevance of recent publications. Accordingly, citation-based comparisons across different time periods should be approached with this limitation in mind, and the apparent underperformance of post-2020 publications in citation metrics is likely, at least in part, a methodological artifact rather than a true indicator of diminished scholarly influence.

Disease-specific analysis demonstrated substantial imbalances in research focus across chronic pediatric conditions. Diabetes mellitus and asthma account for a large proportion of the existing literature, which is consistent with their high prevalence and their well-documented effects of metabolic dysregulation, inhaled medications, and alterations in salivary flow rate on oral health[3,8]. In contrast, conditions such as sickle cell anemia, juvenile idiopathic arthritis, and autism spectrum disorder remain notably underrepresented, despite their well-recognized clinical complexity and potential impact on the delivery of oral health care. This disparity is concerning, as children affected by these conditions often face additional barriers to dental care, including pain crises, functional limitations, behavioral challenges, and increased need for sedation or hospital-based treatment. The relative absence of focused research in these populations suggests not only gaps in scientific knowledge but also potential inequities in the development of evidence-based preventive and therapeutic strategies for vulnerable pediatric groups.

The temporal evolution identified in this study mirrors broader trends in pediatric healthcare. The sharp increase in publications after 2020 likely reflects heightened awareness of healthcare access barriers during the coronavirus disease-2019 pandemic, which disproportionately affected children with chronic diseases requiring hospital-based dental care[12,13]. While this surge signals growing interest in integrated care, sustained efforts will be required to translate emerging evidence into routine clinical practice.

Overall, this analysis highlights both progress and persistent limitations within the domain. Although research activity has increased and patient-centered perspectives are gaining momentum, the lack of integrated frameworks continues to hinder equitable and effective dental care for children with chronic diseases. Addressing these gaps will require interdisciplinary collaboration, targeted research investment in neglected conditions, and policies that support coordinated, preventive, and child-centered oral health care.

Study Limitations

The present study has several limitations that should be acknowledged. The literature search was confined to the WoS database. Although WoS is widely regarded as a rigorous and high-quality source for bibliometric analyses, relying on a single database may not capture the full spectrum of relevant publications indexed exclusively in other repositories such as Scopus or PubMed. Consequently, some studies that could have contributed to the corpus may have been overlooked. Future research is encouraged to replicate or extend this analysis by incorporating multiple databases to enhance the comprehensiveness and generalizability of the findings. Another limitation concerns the temporal coverage of the dataset. Although the analysis encompasses publications up to the year 2026,

it is important to clarify that data were extracted from the WoS database on January 15, 2026. Therefore, only publications indexed up to this specific date were included, and the 2026 data represent a partial rather than a complete annual record. As bibliographic databases may experience indexing delays, some publications from late 2025 and early 2026 might not yet have been indexed at the time of data retrieval. Readers are advised to interpret temporal trend analyses for the most recent period-particularly 2025 and 2026-with this caveat in mind, as the apparent decline or plateau in publication output may partly reflect incomplete indexing rather than a genuine shift in research activity. Future studies updating this analysis after full indexing of 2026 records would provide a more complete picture of recent trends. Methodological limitations inherent to bibliometric and NLP approaches should also be acknowledged. Analytical outcomes are influenced by text preprocessing choices, clustering assumptions, and topic modeling parameters, which may affect the interpretation of thematic structures. Topic identification relied on abstract-level data and keyword-based disease classification, which may have resulted in missed studies or misclassification due to variations in terminology. Moreover, co-occurrence network analysis reflects term proximity rather than causal or clinical relationships.

CONCLUSION

Finally, this analysis describes patterns within the published literature and does not assess study quality, clinical effectiveness, or patient outcomes. As such, the findings provide insight into research trends and gaps rather than providing direct guidance for clinical decision-making.

This NLP-based analysis provides a comprehensive overview of the intellectual structure and thematic evolution of dental caries management research in children with chronic diseases. The findings reveal significant disparities in disease representation, methodological limitations, and insufficient integration of oral health within chronic pediatric care. Addressing these gaps will require more disease-specific, longitudinal, and multidisciplinary research approaches, alongside stronger alignment between dental and medical services. Improving preventive dentistry and effective caries management should be considered an essential component of comprehensive care for children living with chronic diseases.

Ethics

Ethics Committee Approval: This study was based exclusively on publicly available publication data retrieved from bibliometric databases. The analysis did not involve any human subjects, primary data collection, or interventions. Since all data used in this research were already in the public domain and freely accessible, and no identifiable personal information or sensitive data were processed, ethical approval from an institutional review board was not required for this bibliometric analysis.

Informed Consent: Not applicable.

Declaration of AI Use: All authors made substantial contributions to the manuscript. AI assistance was used for language editing, including OpenAI's ChatGPT for English translation and Anthropic's Claude AI for linguistic review and professional refinement. The authors retain full responsibility for the content.

Footnotes

Author Contributions

Concept: H.D., E.E.K., S.M., Design: H.D., E.E.K., S.M., Data Collection or Processing: H.D., Analysis or Interpretation: H.D., E.E.K., Literature Search: H.D., E.E.K., S.M., Writing: H.D., E.E.K., S.M.

Conflict of Interest: The authors declare no conflicts of interest.

Financial Disclosure: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

1. Chen X, Jia L, Wang Q, Wang JJ, Tian Y, Zhang Z, et al. Global, regional, and national burden of caries in primary teeth from 1990 to 2021: results from the global burden of disease study 2021. BMC Oral Health. 2025;25(1):1381. doi: 10.1186/s12903-025-06691-y.
2. Akhter R, Hassan NMM, Martin EF, Muhit M, Smithers-Sheedy H, Badawi N, et al. Caries experience and oral health-related quality of life (OHRQoL) of children and adolescents with cerebral palsy in a low-resource setting. BMC Oral Health. 2019;19(1):15. doi: 10.1186/s12903-018-0704-2.
3. Potgieter N, Pereira V, Elias R, Charone S, Groisman S. Acceptance of the use of silver fluoride among Brazilian parents of children with special health care needs. Front Oral Health. 2024;5:1377949. doi: 10.3389/froh.2024.1377949.
4. Saraç F, Derelioğlu SŞ. Assessment of growth and development, dental caries, and dental development in children with congenital heart disease. J Clin Pediatr Dent. 2025;49(3). doi: 10.22514/jocpd.2025.056.
5. Thikkurissy S, Lal S. Oral health burden in children with systemic diseases. Dent Clin North Am. 2009;53(2):351-7, xi. doi: 10.1016/j.cden.2008.12.004.
6. Chin D, Ramalingam L, Harrison J, Silva M. Oral health and related outcomes in children and adolescents with cystic fibrosis: a scoping review. Eur Arch Paediatr Dent. 2024;25(4):455-69. doi: 10.1007/s40368-024-00885-8.

7. Kodaman Dokumacıgil N, Sezer B, Oktay Ş, Alpay H, Kargül B. Dental caries, oral hygiene and salivary characteristics in children with chronic kidney disease: a case-control study. Pediatr Nephrol. 2025;40(8):2627-37. doi: 10.1007/s00467-025-06730-4.
8. Alrashdi M, Alyahya A. Impact of severity of bronchial asthma on oral health in children. Front Oral Health. 2025;6:1594568. doi: 10.3389/froh.2025.1594568.
9. Atif M, Tewari N, Khan SY, Azam I, Jindal MK, Sharma DS. Impact of oral health conditions on oral health-related quality of life in children with special healthcare needs: a systematic review. J Indian Soc Pedod Prev Dent. 2025;43(4):443-56. doi: 10.4103/jisppd.jisppd_293_25.
10. López-Velasco A, Puche-Torres M, Carrera-Hueso FJ, Silvestre FJ. General anesthesia for oral and dental care in paediatric patients with special needs: a systematic review. J Clin Exp Dent. 2021;13(3):e303-12. doi: 10.4317/jced.57852.
11. Gómez-Ríos I, Pérez-Silva A, Serna-Muñoz C, Ibáñez-López FJ, Periago-Bayonas PM, Ortiz-Ruiz AJ. Deep sedation for dental care management in healthy and special health care needs children: a retrospective study. Int J Environ Res Public Health. 2023;20(4):3435. doi: 10.3390/ijerph20043435.
12. Karaduran B, Kumaş Solak S, Koruyucu M. General anesthesia in pediatric dentistry. Selcuk Dent J. 2022;9(2):696-705. doi: 10.15311/selcukdentj.961790.
13. Doğan S, Günay Canpolat D. Potential complications and preventive measures in pediatric dental treatments performed under general anesthesia and sedation. J Med Dent Invest. 2025. doi: 10.5577/jomdi.e250073.
14. Tasdemir T, Erbas Unverdi G, Ballikaya E, Aypar E, Aykan HH, Karagoz T, et al. Evaluation of oral health status and treatment needs of children with congenital and acquired heart disease. J Clin Med. 2024;13(14):4060. doi: 10.3390/jcm13144060.
15. Vajjala S, Majumder B, Gupta A, Surana H. Practical natural language processing: a comprehensive guide to building real-world NLP systems. Sebastopol (CA): O'Reilly Media; 2020.
16. Chen X, Xie H, Wang FL, Liu Z, Xu J, Hao T. A bibliometric analysis of natural language processing in medical research. BMC Med Inform Decis Mak. 2018;18(Suppl 1):14. doi: 10.1186/s12911-018-0594-x.
17. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: 10.1136/bmj.n71.
18. Miller GA. WordNet: a lexical database for English. Commun ACM. 1995;38(11):39-41. doi: 10.1145/219717.219748.
19. Bird S, Klein E, Loper E. Natural language processing with Python. Sebastopol (CA): O'Reilly Media; 2009.
20. Maier D, Waldherr A, Miltner P, Wiedemann G, Niekler A, Keinert A, et al. Applying LDA topic modeling in communication research: toward a valid and reliable methodology. Commun Methods Meas. 2018;12(2-3):93-118. doi: 10.1080/19312458.2018.1430754.
21. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space [preprint]. arXiv. 2013;arXiv:1301.3781. doi: 10.48550/arXiv.1301.3781.
22. Kusner MJ, Sun Y, Kolkin NI, Weinberger KQ. From word embeddings to document distances. In: Bach F, Blei D, editors. Proceedings of the 32nd International Conference on Machine Learning; 2015 Jul 7-9; Lille, France. PMLR; 2015. p. 957-66. Available from: https://proceedings.mlr.press/v37/kusnerb15.html
23. Ketchen DJ Jr, Shook CL. The application of cluster analysis in strategic management research: an analysis and critique. Strateg Manage J. 1996;17(6):441-58. doi: 10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G.
24. Lloyd SP. Least squares quantization in PCM. IEEE Trans Inf Theory. 1982;28(2):129-37. doi: 10.1109/TIT.1982.1056489.
25. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res. 2003;3:993-1022. doi: 10.1162/jmlr.2003.3.4-5.993
26. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. In: Proceedings of the 8th ACM International Conference on Web Search and Data Mining; 2015 Feb 2-6; Shanghai, China. New York: Association for Computing Machinery; 2015. p. 399-408. doi: 10.1145/2684822.2685324.
27. Salton G, Wong A, Yang CS. A vector space model for automatic indexing. Commun ACM. 1975;18(11):613-20. doi: 10.1145/361219.361220.
28. Waskom ML. Seaborn: statistical data visualization. J Open Source Softw. 2021;6(60):3021. doi: 10.21105/joss.03021.
29. Skeppstedt M, Ahltorp M, Kucher K, Lindström M. From word clouds to word rain: revisiting the classic word cloud to visualize climate change texts. Inf Vis. 2024;23(3):217-38. doi: 10.1177/14738716241236188.
30. Nadeau D, Sekine S. A survey of named entity recognition and classification. Lingvisticae Investig. 2007;30(1):3-26. doi: 10.1075/li.30.1.03nad.
31. Wasserman S, Faust K. Social network analysis: methods and applications. Cambridge: Cambridge University Press; 1994. doi: 10.1017/CBO9780511815478.
32. Hunter JD. Matplotlib: a 2D graphics environment. Comput Sci Eng. 2007;9(3):90-5. doi: 10.1109/MCSE.2007.55.
33. Google. Google Colaboratory [cloud computing platform]. Available from: https://colab.research.google.com
34. Python Software Foundation. Python [software]. 2023. Available from: https://www.python.org/
35. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585(7825):357-62. doi: 10.1038/s41586-020-2649-2.
36. McKinney W. Data structures for statistical computing in Python. In: Proceedings of the 9th Python in Science Conference; 2010 Jun 28-Jul 3; Austin (TX). p. 51-6. doi: 10.25080/Majora-92bf1922-00a.
37. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825-30. Available from: https://jmlr.org/papers/v12/pedregosa11a.html.
38. Seabold S, Perktold J. Statsmodels: econometric and statistical modeling with Python. In: Proceedings of the 9th Python in Science Conference; 2010 Jun 28-Jul 3; Austin (TX). p. 57-61. doi: 10.25080/Majora-92bf1922-011.
39. OpenAI. ChatGPT [large language model]. San Francisco (CA): OpenAI; 2025. Available from: https://chatgpt.com/.
40. Anthropic. Claude [large language model]. San Francisco (CA): Anthropic; 2025. Available from: https://claude.ai/.